Toward an evolutionary perspective on conceptual representation: Species-specific calls activate visual and affective processing systems in the macaque

Ricardo Gil-da-Costa*[†‡§], Allen Braun*, Marco Lopes[‡], Marc D. Hauser[‡], Richard E. Carson[¶], Peter Herscovitch[¶], and Alex Martin[†‖]

*National Institute on Deafness and Other Communication Disorders, [†]National Institute of Mental Health, and [¶]Positron Emission Tomography Department, Clinical Center, National Institutes of Health, Bethesda, MD 20892; [‡]Department of Psychology, Harvard University, Cambridge, MA 02138; and [§]Program in Biology and Medicine, Gulbenkian Science Institute, 2780 Oeiras, Portugal

Non-human primates produce a diverse repertoire of species-specific calls and have rich conceptual systems. Some of their calls are designed to convey information about concepts such as predators, food, and social relationships, as well as the affective state of the caller. Little is known about the neural architecture of these calls, and much of what we do know is based on single-cell physiology from anesthetized subjects. By using positron emission tomography in awake rhesus macaques, we found that conspecific vocalizations elicited activity in higher-order visual areas, including regions in the temporal lobe associated with the visual perception of object form (TE/TEO) and motion (superior temporal sulcus) and storing visual object information into long-term memory (TE), as well as in limbic (the amygdala and hippocampus) and paralimbic regions (ventromedial prefrontal cortex) associated with the interpretation and memory-encoding of highly salient and affective material. This neural circuitry strongly corresponds to the network shown to support representation of conspecifics and affective information in humans. These findings shed light on the evolutionary precursors of conceptual representation in humans, suggesting that monkeys and humans have a common neural substrate for representing object concepts.

auditory | brain | evolution | vocalizations | concepts

Studies of the evolution of animal signaling systems reveal specializations of the peripheral and central nervous systems for producing and perceiving signals linked to survival and reproduction. These signals can be represented, both within and between species, in various modalities, including visual, auditory, olfactory, and tactile, and are often designed to convey considerable information about the signaler and its socioecological context. In particular, non-human primates have evolved complex auditory communication systems that can convey information about a variety of objects and events, such as individual identity (1), motivational state (2), reproductive status (3), body size (4), types of food (5), and predators (6). It has been argued that many of these calls are functionally linked to rich conceptual representations (5–8).

In humans, evidence from behavioral (9), neuropsychological (10), and functional brain-imaging (11) studies suggests that conceptual representations are directly grounded in perception, action, and emotion. For example, functional brain-imaging studies have shown that tasks probing knowledge of animate things and social interactions activate a well defined network that includes regions in the posterior cortex associated with perceiving their visual form (including the fusiform face area) (12) and biological motion [superior temporal sulcus (STS)] (13), as well as limbic and paralimbic cortical areas involved in perceiving and modulating affect (especially the amygdala and medial prefrontal cortex) (11, 14). Neural responses in these regions are elicited by a variety of input modalities, such as visual (12, 15) or auditory (16, 17), and by imagining (18) and thinking about objects and social interactions (19), providing additional support for the idea that the response of these areas is associated with object concept type, not the physical features of the stimulus.

At present, we have little understanding of how the circuitry underlying human conceptual representation evolved and whether we share all or some parts of this circuitry with other animals. Previous studies (mostly behavioral) in non-human primates have reported the capacity for conceptual distinctions in various domains, such as food (20), number (21, 22), and tools (23). Furthermore, studies of visual categorization in non-human primates show both behavioral and neural parallels with studies of humans, including the perception of faces, facial expressions, biological motion, and socially relevant action (13, 24–26). What is unclear, however, is the extent to which these similarities transcend the visual modality, which would provide evidence that, like humans, non-human primates also have more abstract representations of some of these concepts (27–29).

To begin addressing some of these issues, we investigated the neural substrate underlying perception of species-specific calls in rhesus monkeys. We focused on rhesus macaque vocalizations because both behavioral (1, 2, 5) and neurophysiological (30) data suggest that they can be classified on the basis of their meaning, thus transcending their raw acoustic morphology. Previous neurophysiological studies in the macaque that focused on discrete brain areas found evidence for specialized processing of species-specific calls in the auditory (31) and prefrontal (32) cortices. In addition, Poremba *et al.* (33), using flourodeoxyglucose positron emission tomography (PET), reported a hemispheric asymmetry in auditory cortices for species-specific vocalizations. These studies provide some evidence of specificity, opening the door to an investigation of the mapping among call type, conceptual representation, and affective substrate.

In our study, we used the [^{15}O]water PET technique, which allowed us to evaluate processing of species-specific vocalizations in fully awake animals. We were particularly interested in responses elicited outside of the auditory cortices, because they should reflect more conceptual and amodal aspects of these calls.

Materials and Methods

Subjects and Stimuli. Three adult rhesus macaques, *Macaca mulatta* (one male and two females), served as subjects. All procedures and animal care were conducted in accordance with the National Institutes of Health Guide for the Care and Use of

---

Fig. 1. All acoustic stimuli elicited bilateral activation of the auditory cortex. (*A–C*) Waveforms for the three categories of acoustic stimuli used in this study: coos (*A*), screams (*B*), and nonbiological sounds (*C*). Graphs illustrate the wide variation of acoustic morphology among conditions and how the nonbiological sound exemplars were selected to be within the same duration range of the species-specific calls. (*D*) Axial section shows significant bilateral STG activity to all acoustic stimuli in a single monkey.

Laboratory Animals. During each scanning session, three types of auditory stimuli were presented ("coos," "screams," and "nonbiological sounds"; 36 different exemplars of each). We selected these rhesus calls in particular for two reasons. First, coos and screams are acoustically distinctive and associated with different social contexts and states of arousal, and they evoke different responses in listeners. Second, given the requirement within PET studies of contrasting effects across conditions, we wanted to make sure that we could present a sufficiently large and varied sample of exemplars to represent each call type. Of the recordings available (see below), our richest sample of calls was from the categories of coos and screams.

The three auditory stimulus types consisted of the following. (*i*) Coos (Fig. 1*A*): rhesus macaque affiliative vocalizations given in a wide variety of social contexts, including friendly approaches, group movement, calling to individuals out of visual range, and approaching food (usually low-quality, common food). Descriptive statistics on the sample of coos presented are as follows: mean duration = 0.457 sec; range = 0.233–0.888 sec. (*ii*) Screams (Fig. 1*B*): submissive rhesus macaque vocalizations given by subordinates after a threat or attack by a dominant, as well as during the recruitment of coalition support. Descriptive statistics on the sample of screams presented are as follows: mean duration = 0.547 sec; range = 0.122–1.320 sec. Only tonal/arched screams were used (34). (*iii*) Nonbiological sounds (Fig. 1*C*): sounds produced by nonbiological sources, including musical instruments, environmental sounds, and computer-synthesized noises. The mean duration of these sounds was tailored to match the duration range of the coos and screams. Descriptive statistics on the sample of nonbiological sounds presented are as follows: mean duration = 0.498 sec; range = 0.123–1.320 sec. (See *Supporting Text*, which is published as supporting information on the PNAS web site, for further details on animal training, scanner setup, and stimuli sets.)

Based on findings from a pilot study, each acoustic stimulus type was presented during separate scans for a duration of 30 sec, with an interstimulus interval between exemplars that varied randomly from 0.5 to 1.25 sec (see *Supporting Text*). In each session, three to four scans of each of these acoustic conditions were obtained. Sessions occurred on separate days. Each animal participated in 5–6 PET scan sessions over ≈2 weeks, to obtain 16 scans per condition. The presentation order of the exemplars during each scan, as well as the order of the scan conditions, was pseudorandomized to ensure novelty and avoid habituation. None of the acoustic stimuli had previously been presented during training (see *Supporting Text*).

Image Acquisition. *MRI images.* We acquired MRI scans by using a 1.5 T Signa scanner (General Electric) while animals were anesthetized. T1-weighted magnetic resonance images were obtained by using a 3D volume spoiled gradient recall (SPGR) pulse sequence (echo time = 6 ms, repetition time = 25 ms, flip angle = 30°, field of view = 11 cm, slice thickness = 1 mm).
PET images. The lights were dimmed during the scanning sessions. Animals were placed in the scanner gantry while sitting in a specially designed chair with the head restrained to restrict movement (see *Supporting Text*). Stimuli were presented through an audio speaker positioned ≈4 feet in front of the fully awake animal. The speaker, as well as all other objects in the room, were hidden from view by a canopy. PET images were acquired in 2D mode with a GE Advance scanner (General Electric). For each scan, stimulus presentation was initiated simultaneously with the beginning of i.v. injection of 50 mCi (1 Ci = 37 GBq) of [^{15}O]water diluted in 13 ml of saline. Injections were delivered over a 20-sec period, with a 5-min period between injections. Scans 60 sec in duration were acquired. Each contained 35 slices of size $2 \times 2 \times 2$ mm; resolution was 6.5 mm full width at half

maximum in the *x*, *y*, and *z* axes. Images were reconstructed by using a transmission scan for attenuation correction.

Data Analysis. PET brain image analyses were performed by using software incorporated in MEDX (Sensor Systems, Sterling, VA). Structural analysis and realignment of the multiple PET images were accomplished by using FLIRT 3.1 (FMRIB's Linear Image Registration Tool, Oxford Center for Functional Magnetic Resonance Imaging, Oxford). Functional statistics (*t*-maps) were calculated by using statistical parametric mapping (SPM99) (Wellcome Department of Cognitive Neurology, London). Separate analyses were performed on the data acquired in each animal. To identify regions selectively activated for species-specific vocalizations (both coos and screams) vs. nonbiological sounds, significant clusters were defined by a conjunction analysis for voxels more active for both coos vs. nonbiological sounds and screams vs. nonbiological sounds ($P < 0.05$). To identify regions selectively activated by screams, significant clusters were defined by voxels more active for screams vs. coos and more active for screams vs. nonbiological sounds ($P < 0.05$). The functional activation *t*-maps were transformed and realigned with FLIRT (intermodality routine and shadow transform functions) by using each animal's MRI image as a reference. PET and high-resolution structural MRI images were fused to determine the anatomical location of the active areas in MRI space (1 mm^3). The anatomy was defined by using ref. 35. The locations of all activations were confirmed by three experts in rhesus macaque neuroanatomy who were blind to the conditions producing the activations (see Acknowledgments).

Data for the group analyses were obtained by extracting regional cerebral blood flow (rCBF) values from each cluster of activation (minimum size = 50 mm^3), yielding 16 independent measurements for each stimulus condition per animal. These data were submitted to a repeated measures ANOVA with condition (coos, screams, and nonbiological sounds) and scans (16 levels) as main factors. For all analyses reported, there was a significant main effect for condition but not for scans or the condition-×-scans interaction. Correlation analyses were performed to examine regional connectivity patterns by using PET images normalized for global blood flow. For each animal, normalized rCBF values in a "seed voxel" of interest were correlated with all other voxels in the image across the 32 scans acquired during presentation of conspecific vocalizations. Pearson product-moment correlation coefficients were transformed to standard scores (Fisher's z prime transformation) to evaluate significance.

Results

All auditory stimuli, taken together, produced strong, bilateral activity in the belt and parabelt regions of the auditory cortex [superior temporal gyrus (STG)] (Fig. 1*D*). Having established that our experimental procedure reliably produced these expected results, we proceeded to identify areas outside of the STG associated with perceiving species-specific vocalizations. For this purpose, we identified regions that showed a significantly enhanced response to each of the species-specific calls vs. nonbiological sounds. This conjunction analysis revealed robust activity in several posterior visual-processing regions extending from early (V2, V3, and V4) to higher-order (TEO and TE) areas in the ventral object processing stream (Fig. 2 *B* and *D*). In addition, presentation of conspecific vocalizations elicited activity in visual-motion processing [including middle temporal (MT) and medial superior temporal (MST)] areas and, more specifically, biological-motion processing (STS) areas (Fig. 2 *C* and *E*).

To evaluate the possibility that activity in visual-processing areas was due to increased eye movements in response to the calls, we looked for activations within the frontal eye fields (FEF) and correlations between activity in the FEF and the rest

of the brain. The correlational analysis was accomplished by using the coordinates of the peak rCBF value found in the FEF as the seed voxel in each animal. Increased FEF during processing of conspecific calls relative to nonbiological sounds was observed in a single animal. In none of the animals was FEF activity significantly correlated with any of the visual-processing regions described above ($P > 0.10$).

Species-specific calls also convey important information about affective content (25). When animals coo, the animal's emotional state is calm, affiliative, or positively excited because of the presence of food (2). In contrast, when an animal screams, it is in an agonistic state of fear or pain after attack from a conspecific (34). To address the representation of affective content, we compared the neural response to the two types of conspecific vocalizations used in our study. Screams are the more salient of the two and are considered to have a negative and stronger affective value than coos.

Consistent with this behavioral distinction, screams yielded heightened activity in the ventral medial prefrontal cortex (VMPFC) area 32 (Fig. 3 *B* and *E*), the amygdala (Fig. 3 *C* and *F*), and the hippocampus (Fig. 3 *D* and *G*) relative to both coos and nonbiological sounds. In each animal, the lateralization of these activations corresponded to those found in areas TE and TEO (compare Figs. 2 and 3).

Discussion

Our results demonstrate that playbacks of rhesus monkey vocalizations trigger neural activity in visual- and affective-processing areas, as well as in auditory-processing areas. Of greatest significance is the fact that visual areas were activated in the absence of differences in visual input between conditions and that the topology of this network mirrors that observed in humans during conceptual processing. Together, these results suggest that rhesus monkey vocalizations are associated with a conceptual system that is amodal or at least multimodal. This neural system may represent the substrate upon which the human semantic system evolved.

Activation of Higher-Order Visual Cortical Areas by Species-Specific Vocalizations. Previous studies in the rhesus macaque documented that many of the regions activated by species-specific vocalizations are necessary for perceiving object form (V4, TEO, and TE) and biological motion (STS) and in storing object information in memory (especially area TE) (26, 36–38). Detailed anatomical studies have established strong connections between theses areas in the macaque brain. Area TEO has strong reciprocal connections to low-order visual areas such as V2, V3, and V4 and also to STS and TE (39). As noted previously, in humans, high-order visual areas are activated by meaningful auditory stimuli, such as words and object-associated sounds (16, 17). Thus, the finding that this ensemble of visual-object processing areas is activated by meaningful acoustic material in macaques indicates a homologous system in non-human primates and humans for representing object information.

An alternative interpretation of our results is that the activation of these visual-processing regions is due to visual search elicited in the animals by species-specific calls. There are two reasons why this explanation is unlikely. First, differential eye movements should result in increased activation of the FEF. This effect was observed only in a single animal and not in the other two. Second, if activation of visual areas were due to increased visual search, then activity in these regions and the FEF should be strongly linked. However, interregional correlational analysis failed to reveal a functional linkage between the FEF and the regions of the visual cortex reported above in any of the three animals. Overall, then, we conclude that visual search and accompanying eye movements did not significantly contribute to the activation of visual-processing regions in response to con-
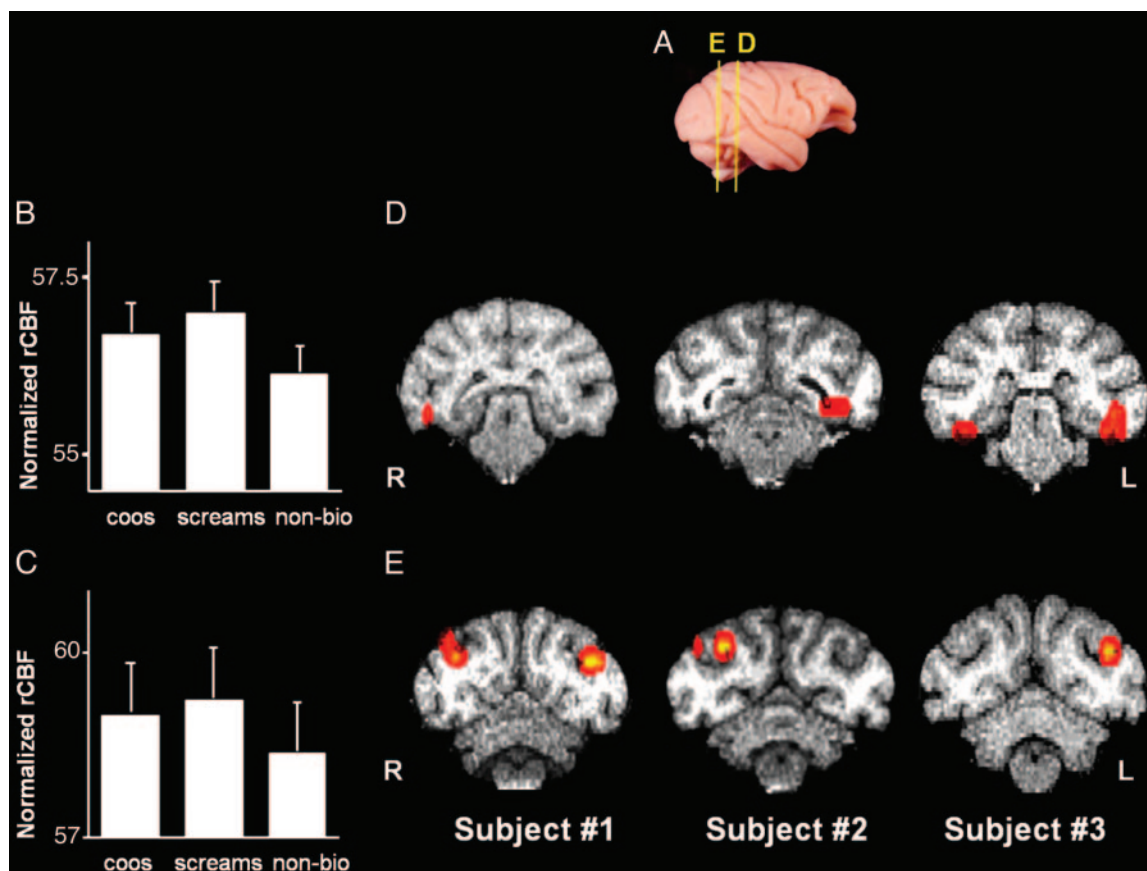
Fig. 2. Selective activation of higher-order visual cortical areas in the rhesus macaque elicited by species-specific vocalizations. (*A*) Lateral view of a rhesus monkey brain illustrates approximate locations of the activations shown in the coronal sections in *D* and *E*. (*B* and *C*) Mean (±SEM) normalized rCBF for the activations in TE/TEO (*B*) and MT/MST/STS (*C*). Group analysis using a repeated measures ANOVA of bilateral rCBF values with planned comparisons revealed that coos and screams did not differ from each other; both yielded greater activation than nonbiological sounds in TE/TEO (*B*) ($F = 8.902$; $P = 0.05$ corrected), and MT/MST/STS (*C*) ($F = 22.783$; $P < 0.005$ corrected). (*D* and *E*) Coronal slices illustrate the location of enhanced activity for conspecific vocalizations (coos and screams) vs. nonbiological sounds in TE/TEO (*D*) and MT/MST/STS (*E*) in each animal. R, right hemisphere; L, left hemisphere.

specific vocalizations. Rather, we suggest that these activations reflect retrieval of stored visual information associated with these calls: specifically, information about socially relevant objects and actions.

An unresolved issue is how the auditory call information reaches these visual-processing regions. Metabolic tracing studies have shown that auditory stimuli activate a large expanse of the cortex, including the entire extent of the STG, as well as regions of the inferior parietal and prefrontal cortices, but not visual-processing areas within the ventral stream observed in our study (V2, V3, TEO, and TE) (40). Studies of humans and macaques, however, provide converging evidence for the integration of auditory and visual information in a posterior region of the STG. In humans, functional imaging studies have identified a region in the posterior superior temporal cortex that integrates complex auditory and visual information (41, 42) that is close to but distinct from the region of STS specialized for processing biological motion (41). The likely homologue of this region in the macaque [area TPO or superior temporal polysensory area (STP)] receives substantial projections from auditory and visual-association cortices (43) and responds to complex auditory and visual stimuli (44). Moreover, Poremba *et al.* (33) reported that species-specific calls selectively activate the TPO in the macaque, and we have observed a similar selective response in our study. Thus, this caudal region of the STG may provide the neural substrate whereby the auditory calls could

activate visual cortices, as well as the subsequent integration of information from these two modalities.

Activation of Limbic and Paralimbic Cortical Areas by Rhesus Macaque Screams. Area 32 in the VMPFC, the amygdala, and the hippocampus were all selectively activated by screams, with each region showing an enhanced response to these affectively loaded calls relative to coos and nonbiological sounds. Although little is known about the function of area 32 in the macaque, electrophysiological studies have shown that this region is highly responsive to conspecific vocalizations, especially signals with high emotional valence (45), and that electrical stimulation of this region elicits species-specific vocalizations in some nonhuman primates (46). In addition, the anatomical connectivity of this region makes it particularly well suited to integrate sensory and limbic information. Area 32 receives most of its sensory input from the anterior auditory-association cortices and also receives some input from early and higher-order (TE) visual areas (47). Moreover, this region has strong connections to the amygdala and hippocampus (47, 48).

In both human and non-human primates, the amygdala and hippocampus play a well documented and critical role in the formation of emotional memories (49, 50). Thus, the heightened response that we observed to screams in the amygdala and hippocampus likely reflected their greater salience and affective significance relative to the other stimuli presented in our study (6). In humans, the VMPFC has been linked to the regulation of
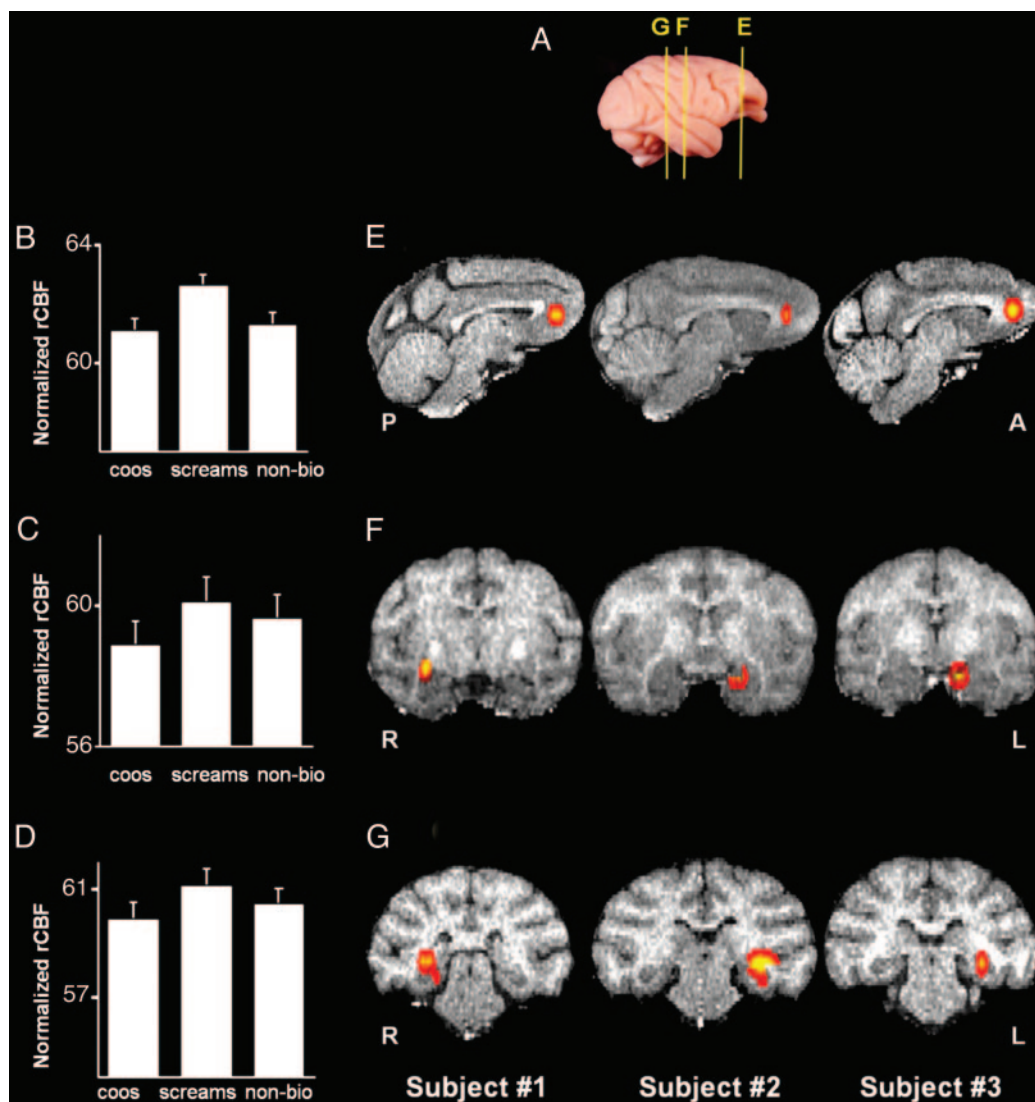
Fig. 3. Selective activation of limbic and paralimbic cortical areas in the rhesus macaque elicited by rhesus screams. (*A*) Lateral view of a rhesus monkey brain illustrates approximate locations of the activations shown in the sagittal sections in *E* and the coronal sections in *F* and *G*. (*B–D*) Mean (±SEM) rCBF for each acoustic condition for activations in VMPFC area 32 (*B*), amygdala (*C*), and hippocampus (*D*). Group analysis using a repeated measures ANOVA with planned comparisons revealed that screams yielded significantly greater activation than both nonbiological sounds and coos in VMPFC area 32 (*B*) ($F = 18.825$; $P < 0.05$ corrected), the amygdala (*C*) ($F = 27.095$; $P < 0.005$ corrected), and the hippocampus (*D*) ($F = 30.263$; $P < 0.005$ corrected). (*E–G*) Sagittal (*E*) and coronal (*F* and *G*) slices illustrate the location of enhanced activity for screams vs. coos and screams vs. nonbiological sounds in VMPFC area 32 (*E*), amygdala (*F*), and hippocampus (*G*) in each animal. P, posterior; A, anterior.

mood (51) and the representation of knowledge of conspecifics (52). Taken together, these data are consistent with the view that the amygdala, hippocampus, and VMPFC are central nodes in a neural network crucial for encoding and remembering emotionally salient material, regulating affective responses, and representing conspecifics in both humans and macaques.

The design of our study closely paralleled studies in humans that have shown enhanced responses in specific cortical regions to "meaningful" material (object pictures, written and spoken words, and object-associated sounds) relative to "meaningless" stimuli (nonsense objects, words, and sounds). The problem, of course, is that for the macaque we cannot assume that the calls are meaningful in the same sense that pictures and words are meaningful to humans, which strongly constrains the interpretations of the activations we have observed. One possibility is that the activations simply reflect a heightened neural response to highly familiar sounds (calls) relative to less familiar stimuli

(nonbiological sounds). However, we believe this possibility is unlikely for several reasons. First, there is no reason to expect that familiar acoustic stimuli, per se, would engage visual-processing regions associated with object perception and memory. Second, although it could be argued that the calls elicit an enhanced state of vigilance and increased eye movements as the animals searched for the source of the calls, the lack of correlated activity in the FEF and posterior visual-processing regions strongly argued against this type of interpretation. Finally, a simple familiarity response could not account for the differential response we observed between one type of call (screams) and another (coos). Not only did screams produce greater responses in certain regions of the brain, but the locations of these regions precisely paralleled the pattern of response seen in the human brain (53). Humans listening to words that convey emotional states and macaques listening to emotional calls (screams) both show increased neural activity in the amygdala and hippocampus

relative to meaningful but less affectively loaded acoustic stimuli (neutral words for humans and coos for the macaque). Thus, we argue that it is unlikely that our findings can be explained as a nonspecific response to familiar items. Nevertheless, several alternative interpretations remain. For example, the distinctions we observed may reflect neural systems for discriminating between conspecific vs. heterospecific vocalizations, biological vs. nonbiological materials, or any meaningful vs. nonmeaningful stimulus. Distinguishing between these alternatives will require more detailed investigations to test whether differential responses can be observed within the system described in this report for stimuli believed to reflect finer distinctions of meaning. Accomplishing that goal will require presentation of a wider range of species-specific vocalizations (e.g., calls presumed to denote foods, conspecific affiliations, threats, etc.), heterospecific vocalizations, and nonmeaningful stimuli coupled with visual representations of these categories and direct monitoring of eye movements and other physiological parameters (e.g., skin conductance and heart rate).

Our findings, along with other recent investigations (32, 33, 54), suggest that the macaque is an important model system for exploring the evolutionary precursors of conceptual representation in humans. Behavioral studies have demonstrated that macaques can match species-specific vocalizations with the corresponding facial expression, illustrating multimodal perception and auditory–visual integration in a non-human vocal communication system (54). On the neural level, it has been demonstrated that macaque species-specific vocalizations activate specific regions of the STG (31, 33) and the prefrontal cortex (32, 45). Our study extends this knowledge by showing that species-specific calls also activate a system known to be associated with the visual representation of objects and their affective properties. This system may have served as an important substrate for the subsequent evolution of language in humans.

1. Rendall, D., Rodman, P. S. & Emond, R. E. (1996) *Anim. Behav.* **51,** 1007–1015.
2. Hauser, M. D. (1991) *Ethology* **89,** 29–46.
3. Semple, S. & McComb, K. (2000) *Proc. R. Soc. London Ser. B* **267,** 707–712.
4. Fitch, W. T. (1997) *J. Acoust. Soc. Am.* **102,** 1213–1222.
5. Hauser, M. D. & Marler, P. (1993) *Behav. Ecol.* **4,** 194–205.
6. Seyfarth, R. M., Cheney, D. L. & Marler, P. (1980) *Science* **210,** 801–803.
7. Zuberbuhler, K., Cheney, D. L. & Seyfarth, R. M. (1999) *J. Comp. Psychol.* **113,** 33–42.
8. Gil-da-Costa, R., Palleroni, A., Hauser, M. D., Touchton, J. & Kelley, J. P. (2003) *Proc. R. Soc. London Ser. B* **270,** 605–610.
9. Barsalou, L. W. (1999) *Behav. Brain Sci.* **22,** 577–660.
10. Warrington, E. K. & Shallice, T. (1984) *Brain* **107,** 829–854.
11. Martin, A. & Chao, L. L. (2001) *Curr. Opin. Neurobiol.* **11,** 194–201.
12. Kanwisher, N., McDermott, J. & Chun, M. M. (1997) *J. Neurosci.* **17,** 4302–4311.
13. Puce, A. & Perrett, D. (2003) *Philos. Trans. R. Soc. London B* **358,** 435–445.
14. Adolphs, R. (2001) *Curr. Opin. Neurobiol.* **11,** 231–239.
15. Chao, L. L., Haxby, J. V. & Martin, A. (1999) *Nat. Neurosci.* **2,** 913–919.
16. Chee, M. W. L., O'Craven, K. M., Bergida, R., Rosen, B. R. & Savoy, R. L. (1999) *Hum. Brain Mapp.* **7,** 15–28.
17. Tranel, D., Damasio, H., Eichhorn, G. R., Grabowski, T., Ponto, L. L. & Hichwa, R. D. (2003) *Neuropsychologia* **41,** 847–854.
18. O'Craven, K. M. & Kanwisher, N. (2000) *J. Cognit. Neurosci.* **12,** 1013–1023.
19. Martin, A. & Weisberg, J. (2003) *Neuropsychology* **20,** 575–588.
20. Santos, L. R., Hauser, M. D. & Spelke, E. S. (2001) *Cognition* **82,** 127–155.
21. Dehaene, S., Molko, N., Cohen, L. & Wilson, A. J. (2004) *Curr. Opin. Neurobiol.* **14,** 218–224.
22. Hauser, M. D., Carey, S. & Hauser, L. B. (2000) *Proc. R. Soc. London Ser. B* **267,** 829–833.
23. Santos, L. R., Miller, C. T. & Hauser, M. D. (2003) *Anim. Cogn.* **6,** 269–281.
24. Freedman, D. J., Riesenhuber, M., Poggio, T. & Miller E. K. (2001) *Science* **291,** 312–316.
25. Hauser, M. D. & Akre, K. (2001) *Anim. Behav.* **61,** 391–400.
26. Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B. & Tootell, R. B. H. (2003) *Nat. Neurosci.* **6,** 989–995.
27. Martin, A. (1998) in *The Origin and Diversification of Language*, eds. Jablonski, N. G. & Aiello, L. C. (California Academy of Sciences, San Francisco), pp. 69–88.
28. Hauser, M. D., Chomsky, N. & Fitch, W. (2002) *Science,* **298,** 1569–1579.
29. Seyfarth, R. M. & Cheney, D. L. (2003) *Annu. Rev. Psychol.* **54,** 145–173.
30. Ghazanfar, A. A. & Hauser, M. D. (2001) *Curr. Opin. Neurobiol.* **11,** 712–720.
31. Rauschecker, J. P., Tian, B. & Hauser, M. D. (1995) *Science* **268,** 111–114.
32. Romanski, L. M. & Goldman-Rakic, P. S. (2002) *Nat. Neurosci.* **5,** 15–16.
33. Poremba, A., Malloy, M., Saunders, R. C., Carson, R. E., Herscovitch, P. & Mishkin, M. (2004) *Nature* **427,** 448–451.
34. Gouzoules, H. & Gouzoules, S. (2000) *Anim. Behav.* **59,** 501–512.
35. Paxinos, G., Huang, X. & Toga, A. W. (2000) *The Rhesus Monkey Brain in Stereotaxic Coordinates* (Academic, San Diego).
36. Desimone, R., Albright, T. D., Gross, C. G. & Bruce, C. (1984) *J. Neurosci.* **4,** 2051–2062.
37. Logothetis, N. K., Guggenberger, H., Peled, S. & Pauls, J. (1999) *Nat. Neurosci.* **2,** 555–562.
38. Naya, Y., Yoshida, M. & Miyashita, Y. (2001) *Science* **291,** 661–664.
39. Distler, C., Boussaud, D., Desimone, R. & Ungerleider, L. G. (1993) *J. Comp. Neurol.* **334,** 125–150.
40. Poremba, A., Saunders, R. C., Crane, A. M., Cook, M., Sokoloff, L. & Mishkin, M. (2003) *Science* **299,** 568–572.
41. Beauchamp, M. S., Lee, K. E., Argall, B. D. & Martin, A. (2004) *Neuron* **41,** 809–823.
42. Van Atteveldt, N., Formisano, E., Goebel, R. & Blomert, L. (2004) *Neuron* **43,** 271–282.
43. Seltzer, B., Cola, M. G., Gutierrez, C., Massee, M., Weldon, C. & Cusick, C. G. (1996) *J. Comp. Neurol.* **370,** 173–190.
44. Bruce, C., Desimone, R. & Gross, C. G. (1981) *J. Neurophysiol.* **46,** 369–384.
45. Vogt, B. A. & Barbas, H. (1988) in *The Physiological Control of Mammalian Vocalization*, ed. Newman, J. D. (Plenum, New York), pp. 203–225.
46. Robinson, B. W. (1967) *Physiol. Behav.* **2,** 345–354.
47. Barbas, H., Ghashghaei, H. T., Rempel-Clower, N. L. & Xiao, D. (2002) in *Handbook of Neuropsychology*, ed. Grafman, J. (Elsevier Science, Amsterdam), pp. 1–27.
48. Insausti, R. & Muñoz, M. (2001) *Eur. J. Neurosci.* **14,** 435–451.
49. Baxter, M. G. & Murray, E. A. (2002) *Nat. Rev. Neurosci.* **3,** 563–573.
50. Packard, M. G. & Cahill, L. (2001) *Curr. Opin. Neurobiol.* **11,** 752–756.
51. Drevets, W. C., Price, J. L., Simpson, J. R., Jr., Todd, R. D., Reich, T., Vannier, M. & Raichle, M. E. (1997) *Nature* **386,** 824–827.
52. Mitchell, J. P., Heatherton, T. F. & Macrae, C. N. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 15238–15243.
53. Kensinger, E. A. & Corkin, S. (2004) *Proc. Natl. Acad. Sci. USA* **101,** 3310–3315.
54. Ghazanfar, A. A. & Logothetis, N. K. (2003) *Nature* **423,** 937–938.

NEUROSCIENCE